Longitudinal Data from Social Security Records in Austria

Hofer, Helmut; Winter-Ebmer, Rudolf

Schmollers Jahrbuch 123 (2003), 587 – 591
Duncker & Humblot, Berlin

European Data Watch

This section will offer descriptions as well as discussions of data sources that may be of interest to social scientists engaged in empirical research or teaching courses that include empirical investigations performed by students. The purpose is to describe the information in the data source, to give examples of questions tackled with the data and to tell how to access the data for research and teaching. We will start with data from German speaking countries that allow international comparative research. While most of the data will be at the micro level (individuals, households, or firms), more aggregate data and meta data (for regions, industries, or nations) will be included, too. Suggestions for data sources to be described in future columns (or comments on past columns) should be send to: Joachim Wagner, University of Lueneburg, Institute of Economics, Campus 4.210, 21332 Lueneburg, Germany, or e-mailed to ⟨wagner@uni-lueneburg.de⟩.

Longitudinal Data from Social Security Records in Austria

By Helmut Hofer and Rudolf Winter-Ebmer

1. Introduction

Work history data from social security records have become available in recent years in many countries (e.g. in Germany or Belgium). These data share many common features: they come from official sources, are collected for "non-research" purposes, in general they serve as evidence for eligibility of old-age pension or other forms of social security payments. This construction has some distinct advantages and disadvantages: the data are (almost[1]) universal, longitudinal – often they cover several decades for each person, they are reasonably exact in terms of employment spells and wages, because tax payments of firms and benefits of workers hinge on these data. Another advantage arises from the universal coverage. If firm identifiers are available, a crude

[1] In many cases workers with very low wages are not covered in the data set.

employer-employee data set is easily formed: the workforce in a firm is easily identified and thus data about wage structure and worker turnover in firms. More elaborate employer-employee data sets require matching of data, for instance with firm data from national statistics or tax data.

Due to the administrative character of these data, some disadvantages are also common. The prime problem is the lack of socio-economic characteristics of the persons. The data are collected for a very specific administrative purpose; extensions of the data set – over and above the directly necessary information – are in general not possible. Due to the universal coverage, adding additional variables would be very costly to the firms and the public administration.

2. Information in the data

In Austria, the insurance relevant information is gathered by the Austrian Social Security Administration (Hauptverband der österreichischen Sozialversicherungsträger). This data source contains daily data on the labour market status for all workers, except for self-employed workers and part of the civil servants. The data set covers the period from January 1972 until now. Currently the data set contains information of about 8.5 Million persons with 45 Million insurance spells (Versicherungszeiten). For convenience we label the social security records SED-data. The data structure has some useful features. One can investigate the cross-sectional as well as the longitudinal character of the data. There is almost no panel attrition and the data are very reliable, as the firms are legally obliged to report correct figures and the figures are required for collecting social security contributions as well as for paying out old-age pensions. Exceptions are immigrant workers returning to their home country or emigrants from Austria. As all individual data on employment spells have a unique firm identifier it is easy to form an employer-employee data set.

Box 1
Variable list for individuals and firm information

Person number
Sex
Age
Nationality
Monthly Salary
Days working, sick, unemployed
Labour Market Status

Establishment number
Location (Districts (Bezirk)
Industry affiliation (4-digit identifier)
Firm Size (16 classes)

Information available in the SED-data set is very comprehensive with respect to social insurance relevant information (see Box 1). Socio-economic characteristics like age, sex, and nationality are available. It is possible to track the employment (and unemployment) history of the worker on a daily base. However, no information about working-time is registered. In principle, the insurance records allow to distinguish insurance coverage due to gainful employment, self-employment, unemployment and maternity leave. The data contain information about the occupational status (white-collar, blue-collar worker, not tenured civil servant, self-employed, apprentice, etc.). For every employment spell we know the begin date and the end date of the spell, monthly earnings and the employer identification. Wages are top coded because of the social security contribution cap. This applies to less than 10% of the working individuals. Information on registered unemployment spells (with / without unemployment benefits), days on sickness leave and on maternity leave are available.

The database is organised on an individual basis; however, the employers' social security number (Dienstgeberkontonummer) allows linking the persons to establishments. It is not possible, however, to distinguish between the plant and the enterprise level, because some employers report all employees for each plant separately, whereas other employers report all employment in various plants under a single employer social security number. With respect to the establishment information on industry affiliation, firm size, and region is available. Major drawbacks concern the lack of information on working hours, on family affiliation and educational attainment.

3. A look at topics of selected studies with the data

Various studies have been undertaken with the SED-data. We will enumerate some selected topics:

The effects of trade and migration are investigated by regressing industry level information for trade and migration on individual labour market careers (Aiginger et al, 1996, Hofer and Huber 2003, Winter-Ebmer and Zweimüller, 1996). These studies show a relatively weak effect of migration and trade on the labour market chances of Austrian workers.

Hofer and Weber (2002) examined wage mobility for Austria. Issues of mobility are particularly interesting with these data because of the long time periods and the precise transitions. Wage mobility can be calculated via measures based on transitions between quintiles in the wage distribution, and by indices, measuring the extent to which averaging wages over a longer period decreases cross-sectional inequality. This study indicates that wage mobility is very low in Austria in international comparison. Other mobility issues concern unem-

ployment entry: Winter-Ebmer (2003) as well as Zweimüller and Lalive (2003) analyse a quasi-experimental situation, where unemployment benefit entitlement was extended only for workers above age 50. As information on all Austrian workers is available, the researchers were able to focus on very close control groups: those aged 49 versus those aged 50.

Furthermore, other studies investigated issues at the firm level. The SED-data set was used to explore job turnover in Austria (see Hofer et al. 2001, Stiglbauer et al, 2003). These studies decompose the net change in employment between two points in time into gross flows due to entering, growing, shrinking and expanding firms. Winter-Ebmer (2001) used data for one big firm to investigate the returns to a specific manpower training program. In such a case, it is easy to match additional (firm-specific or individual) information to the data.

4. Data access

Employment data are strictly confidential by law. However, statistical offices are allowed to give these data to outside researchers, when the data are manipulated so that individual persons or firms cannot be identified. Under that condition the Social Security Administration offers access to the data via the SED (Sozialversicherungsdaten, Externe Schnittstelle) interface. The researcher hat to sign a contract with the Social Security Administration and she receives the data in electronic form. He must pay a fee for the expenses of the Social Security Administration. It is possible to draw a sample or even to get the universe of all insurance spells.

In general there is the possibility to match the SED-data with information from the employment office (AMS) on an individual base. The AMS-data extend the informational content significantly. It contains information on family status, education and occupation. Note, that this information is available only for persons, who have been registered at the AMS at least once. The researcher needs the permission of the employment office to receive the matched data. This permission can be granted for special research projects.

In principle, it would be possible to receive SED-data for teaching courses as long as the anonymity of the data is secured. However, the data have to be bought for that purpose and the data handling is relatively complex.

5. A look ahead

Some drawbacks of the data set could be overcome by combining further data sources. The statistical office Austria (Statistik Austria) conducts a yearly report of the income of all Austrians (Statistik Austria 2003). For this purpose

they combine the Microcensus, tax data on labour income and the SED-data. This database would allow investigating income above the contribution ceiling, and it would offer information on working time and socio-economic characteristics (like education, occupation, family background). Unfortunately, until now this data set is confidential and not useable for research.

The Social Security Administration plans to extend data coverage to include schooling information and some limited information about working time in the future.

6. References

Aiginger, K. / *Winter-Ebmer,* R. / *Zweimüller,* J. (1996): Eastern European Trade and the Austrian Labor Market, Weltwirtschaftliches Archiv.

Hofer, H. / *Huber,* P. (2003): Wage and Mobility Effects of Trade and Migration on the Austrian Labour Market, Empirica 30, 107 – 125.

Hofer, H. / *Pichelmann,* K. / *Schuh,* U. (2001): Price and Quantity Adjustments in the Austrian Labour Market, Applied Economics 33, 581 – 592.

Hofer, H. / *Weber,* A. (2002): Wage Mobility in Austria 1986 – 1996, Labour Economics 9, 563 – 577.

Statistik Austria (2003): Bericht gemäß Art 1 § 8 Bezügebegrenzungsgesetz, für die Jahre 2000 und 2001, Statistik Austria, Vienna.

Stiglbauer, A. / *Stahl,* F. / *Winter-Ebmer,* R. / *Zweimüller,* J. (2003): Job Creation and Job Destruction in a Regulated Labor Market: The Case of Austria, Empirica 30, 127 – 148.

Winter-Ebmer, R. (2003): Benefit duration and unemployment entry: a quasi-experiment in Austria, European Economic Review 47 (2): 259 – 273.

– (2001): Evaluating an innovative redundancy-retraining project: The Austrian Steel Foundation, mimeo, University of Linz.

Winter-Ebmer, R. / *Zweimüller,* J. (1996): Immigration and the Earnings of Young Native Workers, Oxford Economic Papers.

Zweimüller, J. / *Lalive,* R. (2003): Benefit Entitlement and the Labor Market:Evidence from a Large-Scale Policy Change, in: Agell, J. / Keene, M. / Weichenrieder, A. (eds.): Labor Market Institutions and Public Policy, forthcoming, MIT Press.